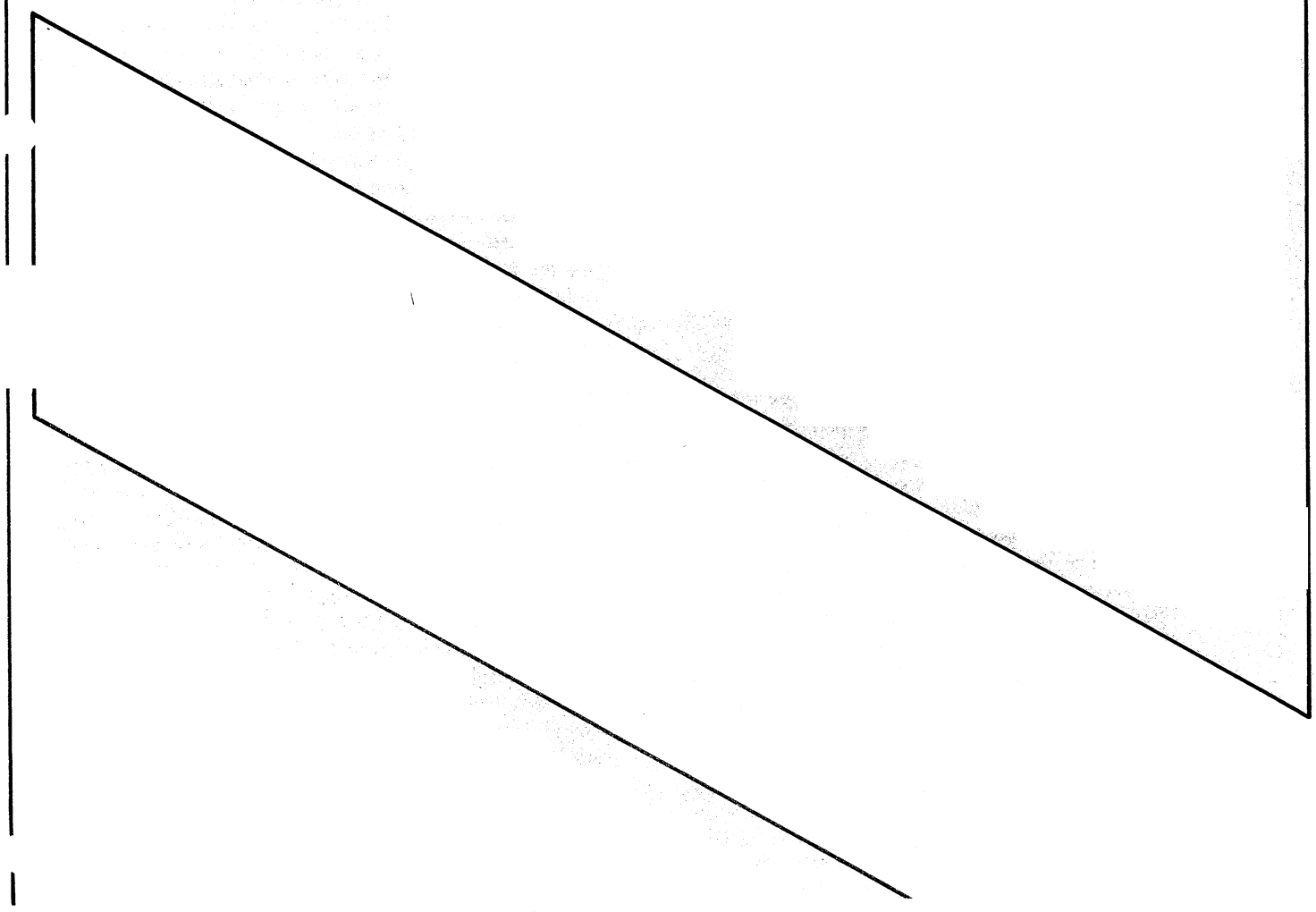
Reprinted from Trends in Food Science and Technology



Molecular modeling in food research: technology and techniques

Thomas F. Kumosinski, Eleanor M. Brown and Harold M. Farrell, Jr

The development of structure–function relationships is essential to the effective use of biotechnology for solving problems in food science and technology. The introduction of more powerful computers, together with the increased availability of efficient molecular modeling software will enable the study of such relationships. In this review, we present the most important techniques used in the modeling of small molecules, proteins, nucleic acids and polysaccharides. A brief description of the application of experimental techniques for structure determination, such as X-ray crystallography, two-dimensional nuclear magnetic resonance spectroscopy, circular dichroism spectroscopy and Fourier transform infrared spectroscopy is provided, with particular emphasis on their relevance to molecular modeling. Various techniques for determining the minimum conformational potential energy, ranging from rigorous molecular orbital methods based on quantum mechanics to approximate molecular mechanics methods based on force-field calculations, are described; the selection of appropriate methods to study particular sizes and types of molecules and molecular motions is also presented.

Biotechnology promises the development, by the new techniques for the genetic engineering of proteins, of products with tailor-made functionalities for food manufacture, and the creation of new co-solutes that may control functionality. However, the historic problem of developing quantitative structure–function relationships still exists. Without knowledge of such relationships, the new techniques are limited in potential and have a low probability of success.

The caseins of bovine milk and their naturally occurring genetic variants provide an illustration of qualitative correlations between protein primary structure and functionality in food systems. For example, cheeses made from milks containing α_{s1}-casein A have a softer texture and body than those made from the more frequently occurring α_{s1}-casein B variant; milks containing the α_{s1}-casein A variant are also more resistant to calcium-induced coagulation. The α_{s1}-casein A variant is the result of the deletion of 13 amino acids (residues 14–26) from the B variant[1]. In β- and κ-caseins, single-site mutations change the chymosin-induced clotting time, the initial step in cheese manufacture[1]. However, the changes in secondary, tertiary and quaternary structure resulting from such mutations have not been obvious. Thus, the mechanism of the functionality changes described above is poorly understood, and the success rate of future deliberately induced mutations cannot be predicted.

In recent years, the development of molecular modeling as a technique for developing or refining three-dimensional molecular structures has resulted in a methodology capable of suggesting a molecular basis for structure–function relationships. At present, the behavior of food proteins, preservatives, emulsifiers and stabilizers can be modeled in a food system. Studies of the structure–function relationships of new peptides, carbohydrates, polysaccharides and small molecules can be used to test their potential effectiveness.

The purpose of this report is to provide a synopsis of the various techniques in molecular modeling and to guide interested food scientists in acquiring the most useful technologies, both software and hardware, for solving their specific problems. A future report in this journal will deal with specific applications of various techniques for solving structure–function relationships.

Techniques for the determination of three-dimensional structure

The traditional method for determining the three-dimensional structures of proteins and smaller molecules is X-ray crystallography. The methodology enables the elucidation of structure with a high degree of precision[2]. In fact, the structure of a small molecule can be determined from the diffraction pattern of its crystals in a relatively short time using so-called 'direct methods' (see Glossary). However, for macromolecules (e.g. proteins and nucleic acids), the solution is not straightforward, and it is generally necessary to prepare several different heavy-atom isomorphous derivatives. The diffraction patterns for the crystal with and without the heavy metal are then obtained and the results are analysed using Patterson synthesis. The limitations of X-ray crystallography for the determination of structure–function relationships in proteins include the requirements that the sample be crystalline and that a suitable heavy metal be incorporated without distorting the crystal.

Mention of brand or firm names does not constitute an endorsement by the US Department of Agriculture over others of a similar nature not mentioned.

Thomas F. Kumosinski, Eleanor M. Brown and **Harold M. Farrell, Jr** are with the US Department of Agriculture, Agricultural Research Service, Eastern Regional Research Center, 600 E. Mermaid Lane, Philadelphia, PA 19118, USA.

Unfortunately, many proteins and other food molecules cannot be crystallized at present and, of course, the functional form of the molecule usually exists in the solution, gel or sol state – not in the form of a crystal.

In contrast to X-ray crystallography, two-dimensional nuclear magnetic resonance (NMR) spectroscopy can measure the structure of a molecule in solution[3]. The methodology is not as precise as X-ray crystallography, is time-consuming and is limited to the study of proteins with less than 100 amino acid residues. Two types of NMR experiment are used: correlation spectroscopy (COSY), which uses the spin–spin coupling between protons to establish band frequencies for specific amino acid residues, and nuclear Overhauser enhancement spectroscopy (NOESY), which measures the 'through space' dipole–dipole coupling between protons. At a given band frequency, the dipole–dipole coupling constants measured by NOESY are directly proportional to the correlation times of the magnetically coupled protons, and are inversely proportional to the cube of the distance between them. Thus, the NOESY data provide a measure of the relative motion of the protons and of the 'through space' distances between them. The calculated 'through space' distances can be used to derive a plausible three-dimensional static structure for a small molecule or protein, provided that there is a large number of well-resolved NMR bands. The larger the number of bands, the greater the precision and accuracy of the resulting structure. Thus, the technique requires a high-field NMR spectrometer; usually a frequency of 500 MHz is needed for acceptable proton resolution and to obtain two-dimensional spectra. In addition, higher molecular weight proteins have longer correlation times, larger nuclear Overhauser effects and larger band widths, all of which decrease the resolution while increasing the number of bands. Hence, the current size limit is 100 amino acid residues.

Techniques for the determination of secondary structure

The most widely used technique for the estimation of the secondary structure of a protein is circular dichroism spectroscopy[4]. The technique measures the dependence on wavelength of ellipticity (the difference in the absorbance of left and right circularly polarized light) of the optically active peptide bonds in the far ultraviolet region, mainly 230–185 nm. Three bands with either positive or negative maximum ellipticities are produced for each of the standard conformational states (α-helix, β-pleated sheet and random coil); one is produced for the $n \rightarrow \pi^*$ transition (energy absorption by a non-bonding electron), and one each for the parallel and perpendicular $\pi \rightarrow \pi^*$ (double-bond absorption) electronic transitions of the electrons of each peptide bond. The spectrum observed for residues in an α-helix conformation is much more intense than those for other conformations; thus, the method is most reliable for quantifying the content of α-helix. The need to have optically clean solutions (any components that scatter light will affect the results) and very accurate determinations of protein concentration complicate the method. However, an advantage of the method is the need for only small volumes of dilute aqueous protein solutions.

On the other hand, Fourier transform infrared (FTIR) spectroscopy experiments can be applied to samples in any state (solid, solution, gel, sol, etc.). However, protein concentrations must be high (20–50 mg/ml), and D_2O rather than H_2O is used for solution-phase experiments. The resolution of the amide I band, which arises from the peptide bonds, can be increased by the use of Fourier transform deconvolution techniques[5]. Individual frequencies within the broad amide I band can be calculated. Since a large signal-to-noise ratio is obtained using FTIR, various mathematical techniques can be used to enhance the spectra. The resulting deconvoluted and enhanced spectra can then be fitted to a sum of individual Gaussian peaks using nonlinear regression analysis to determine the relative contributions of individual bands to the observed broad amide I band. The frequencies of the individual bands can, in turn, be correlated with secondary structural assignments using the theoretical calculations of Krimm and Bandekar[6] in combination with the results of FTIR spectra obtained for proteins for which three-dimensional X-ray crystal structures are available[5]. Thus, the amounts of β-turn, α-helix, β-sheet, extended strand or unordered structure can be estimated. Although this methodology is promising, its precision and accuracy for the determination of turn structures are still being assessed.

Criteria for building molecular structures

Obtaining detailed three-dimensional images of food proteins by traditional methods may be impractical. Many food proteins do not readily form crystals or, at least, crystals suitable for high-resolution X-ray diffraction analysis. Moreover, facilities for structure determination by X-ray crystallography may not be available locally, and finding collaborators who are interested in a food protein may be difficult. However, the construction of a model based on previously determined physicochemical

Glossary

Direct methods: Statistical methods for determining the phases of the diffraction beams in an X-ray diffraction experiment.

Isomorphous derivatives: Another approach to determining the phases in X-ray diffraction experiments that involves replacing an atom in the molecule by a heavy atom, without appreciably distorting the crystal structure.

Patterson synthesis: Yet another approach to the determination of phases in X-ray diffraction experiments. This technique may be used in conjunction with the isomorphous replacement method, and involves the use of the mathematical technique of Fourier synthesis to calculate the electron densities from the diffraction beam intensities.

Nuclear Overhauser effect (NOE): The saturation of the resonance signal that occurs when nuclei are irradiated to the extent that the nuclear energy levels become equally populated. Observation of NOEs can give information about internuclear distances.

Self-consistent field (SCF) orbitals: This refers to the mathematical approach used to calculate the orbitals; initially, a set of orbitals is assumed and the electron–electron repulsion is calculated. The energy obtained is then used to calculate a new set of orbitals. The process is repeated until convergence and self-consistency is achieved for the orbitals.

and spectroscopic characteristics of the protein as well as on the predicted behavior of the amino acid sequence can provide a starting point for the exploration of structure–function relationships.

When deriving molecular models, it is imperative to use a library or a dictionary of geometric parameters to ensure that assigned bond lengths, bond angles, and van der Waals radii are compatible with those determined by X-ray crystallography. The resulting molecular models can then be compared with other experimentally determined structures. All of the major molecular modeling software packages have such libraries or dictionaries. Table 1 lists some of the software packages currently available for molecular modeling. The packages vary in terms of their facilities and possible applications, and the researcher should study such systems in detail to determine the most appropriate system for solving a particular problem that is compatible with the available hardware.

Each of the software packages listed in Table 1 contains libraries or dictionaries that are compatible with geometric parameters derived from X-ray crystallography. They are easily interfaced with the Brookhaven database of protein crystal structures[7] and the Cambridge crystallographic database (Cambridge Crystallographic Data Files, Crystallographic Data Centre, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, UK), which contain most of the available X-ray crystal structures for proteins and small molecules, respectively. Each software package can calculate the potential energy field of a molecule, and incorporates one or more algorithms for minimizing conformational energy and, usually, an algorithm for modeling molecular dynamics. Models of small molecules, polysaccharides, polynucleotides and proteins can be 'built', their conformational energies minimized, and the structures compared with those determined experimentally. Finally, these packages interface easily with several programs, available from the Quantum Chemistry Program Exchange (QCPE) (Quantum Chemistry Program Exchange, Department of Chemistry, Indiana University, Bloomington, IN, USA), that calculate molecular parameters using quantum mechanics.

Molecular orbital calculations

Ab initio self-consistent field (SCF) molecular orbital calculations (based on quantum theory) can be done routinely for small molecules (not proteins)[8]. Such calculations are generally based on some modification of the Hartree–Fock method, which approximates molecular orbitals as a linear combination of hydrogen-like, Slater-type or Gaussian orbitals; this linear combination of atomic orbitals is similar to an infinite series (Taylor series), which expresses a function as a sum of polynomial terms. The Hartree–Fock approximation as generally used is very time consuming, which makes *ab initio* calculations on large molecules impractical. *Ab initio* calculations are usually computed using one of the QCPE programs ('Gaussian 80' or 'Gaussian 86'), and give molecular parameters such as atomic partial charges, dipole moments and quadrupole moments for moderately sized heteroatomic molecules of known crystal structure.

Table 1. Software packages for molecular modeling

Package	Company
Biograph	Bio Design, Inc., Pasadena, CA, USA
Insight, Discover	Biosym Technologies, San Diego, CA, USA
SYBYL	Tripos Associates, Inc., St Louis, MO, USA
CHARMM, Quanta	Polygen, Waltham, MA, USA
Chem-X	Chemical Design Ltd, Oxford, UK
INTERCHEM	Interprobe Chemical Services, Glasgow, UK

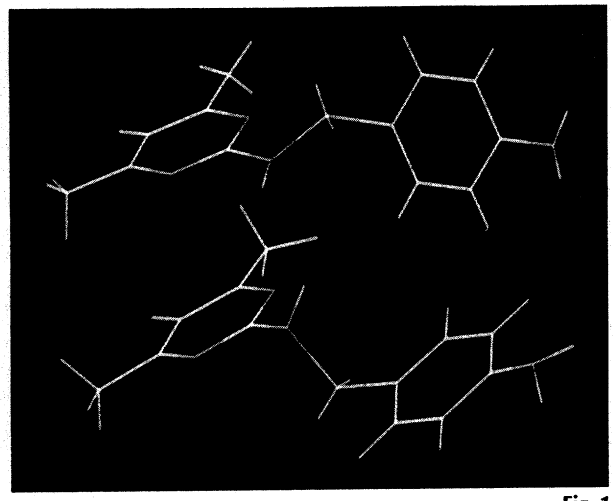


Fig. 1

The structure of sulfamethazine as determined by X-ray crystallography (above), compared with the energy-minimized structure (below). Carbon, white; hydrogen and nitrogen, blue; oxygen, red; sulfur, yellow.

The most useful calculations for identifying the optimum geometry – the structure that has the lowest potential energy – are the semi-empirical techniques[8]. Unlike the entirely theoretically based *ab initio* methods, the semi-empirical methods are based on fitting experimental data to a series of models, and using the parameters obtained to predict the behavior of other molecules. Semi-empirical approaches enable the energy of a large molecule to be computed within a realistic time. Like the *ab initio* approach, these approaches usually start with some form of the general SCF molecular orbital theory; the semi-empirical approaches then make approximations to the quantum mechanics equations in an effort to reduce the number of terms to be computed.

Several semi-empirical programs are available from the QCPE, and differ only in their treatment of the cumbersome coulomb integrals (describing coulombic interactions between the two charge distributions) and exchange integrals; they include the 'Complete neglect of differential overlap' (CNDO), the 'Intermediate neglect of differential overlap' (INDO), the 'Neglect of diatomic differential overlap' (NDDO), the 'Modified neglect of differential overlap' (MNDO) and the 'Modified intermediate neglect of differential overlap' (MINDO and MINDO/3) packages. These programs are useful for finding the lowest conformational energy

state of a molecule by varying atomic distances and/or dihedral angles.

Molecular force-field methods

Due to efficient algorithms, and the lower cost and increased availability of fast computers with large memories, calculations based on quantum mechanics have greatly increased the development of structure–function relationships for small molecules, in which the lowest potential energy and, hence, the most likely geometry are relatively easy to define. However, many problems of biological interest, such as the study of polynucleotide and protein conformation, can still be modeled only by using the most elementary empirical energy functions. Although such models are crude, the approach has been applied successfully to the study of hydrocarbons, oligonucleotides, peptides and amino acids during the past few years.

In the case of molecular mechanics or force-field methods, the molecule is represented by a collection of overlapping balls (the atoms, with given van der Waals radii) connected by springs (which mimic the vibrational character of the bonds). The atoms are assigned certain van der Waals attractive and repulsive forces as well as electrostatic forces, representing non-covalent interactions. Molecular mechanics methods use a combination of potential energy functions to optimize a structure. The three most important requirements for force-field calculations are an equation that calculates energy as a function of molecular geometry, the model parameters (a set of 'best' values for experimentally derived molecular properties), and an algorithm to calculate new atomic coordinates. Possibly, the most theoretically sound force-field package is 'MM2' by Allinger[9], which is available from the QCPE. This package has been successfully used to determine the conformation of minimum potential energy of oligosaccharides and peptides, using an iterative approach[10], but its use is limited to the study of molecules with less than 200 atoms. Other force-field methods, based on more empirical parameters derived from geometric and thermodynamic studies, have been successfully used for large molecules such as proteins, polynucleotides, synthetic polymers and polysaccharides. A recent review[11] examines several programs that use such methods ('ECEPPS'[12], 'CHARMM'[13] and 'AMBER'[14]). Although these software packages are quite sophisticated, they are still under development; at present, they are not able to handle hydrophobic interactions, and they do not take into account any interactions of the molecule with solvent, even in the form of a general term: molecules are studied in complete isolation ('in vacuum'). The size of the molecule for which force-field calculations can be computed is limited by the speed and memory of the computer.

Building sulfamethazine

Sulfamethazine is an important sulfa drug widely used as an antibiotic for treating mastitis in dairy cows[15]. A computer-generated image of its structure, based on the results of X-ray crystallography, is presented in Fig. 1. To test the validity of the molecular modeling algorithms, we used the 'SYBYL' molecular modeling package to build a model of sulfamethazine. The 'Sketch module' facility of the program enables the investigator to 'draw' the basic structure of the molecule from the constituent atoms, from fragments, or by modification of an existing structure. Bond lengths and angles can be adjusted to realistic values. To obtain the molecular structure of lowest potential energy (shown in Fig. 1 for comparison with the structure determined by X-ray crystallography), the energy of the sulfamethazine model was first minimized using a force-field model, which ignored electrostatic interactions. The energy of the resulting structure was further minimized using the MNDO program, to optimize the dihedral angles. The interatomic distances were already compatible with those determined by X-ray crystallography and, thus, were not optimized further. Comparison of the X-ray and computed structures indicates that the rotational angles for the C–N and N–P bonds of the two conformations differ. Since these bonds rotate freely when the molecule is in solution, the two structures may be considered to be in good agreement with each other; the energy difference between the two conformations is low. Atomic partial charges calculated for the X-ray structure using Gaussian 80 were in good agreement with those calculated using MNDO for the

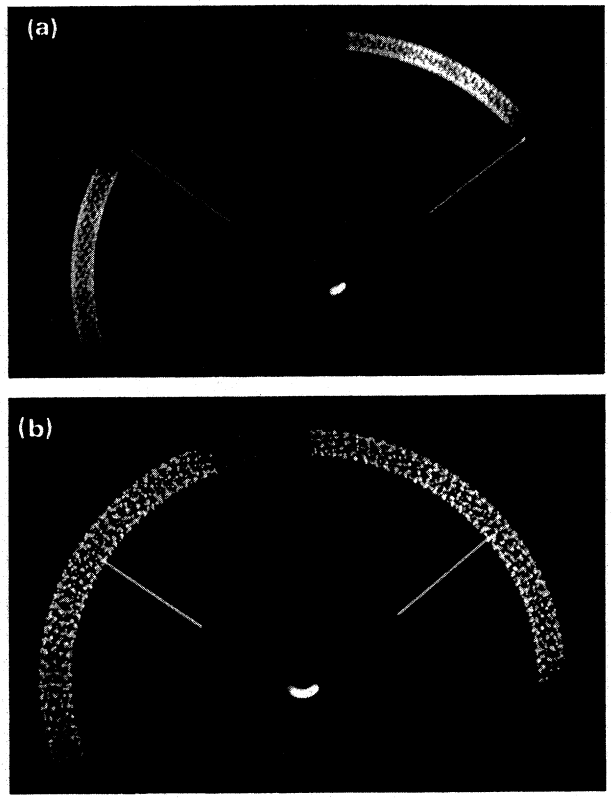


Fig. 2
Sweep graph of the dynamic motion of water, represented by a stick-model 'V', over 25 picoseconds: (a), at 300 K; (b), at 600 K. The white area at the center of each image represents the location of the oxygen atom; arcs describe the locations of the hydrogen atoms during the 25-picosecond observation period. Color coding indicates the proportion of time spent by the hydrogen atoms in a particular location: red, high proportion of time; purple, low proportion of time.

energy-minimized structure. With such techniques it is now possible to examine newly synthesized drugs and other molecules to compare their molecular properties and to determine structure–activity relationships.

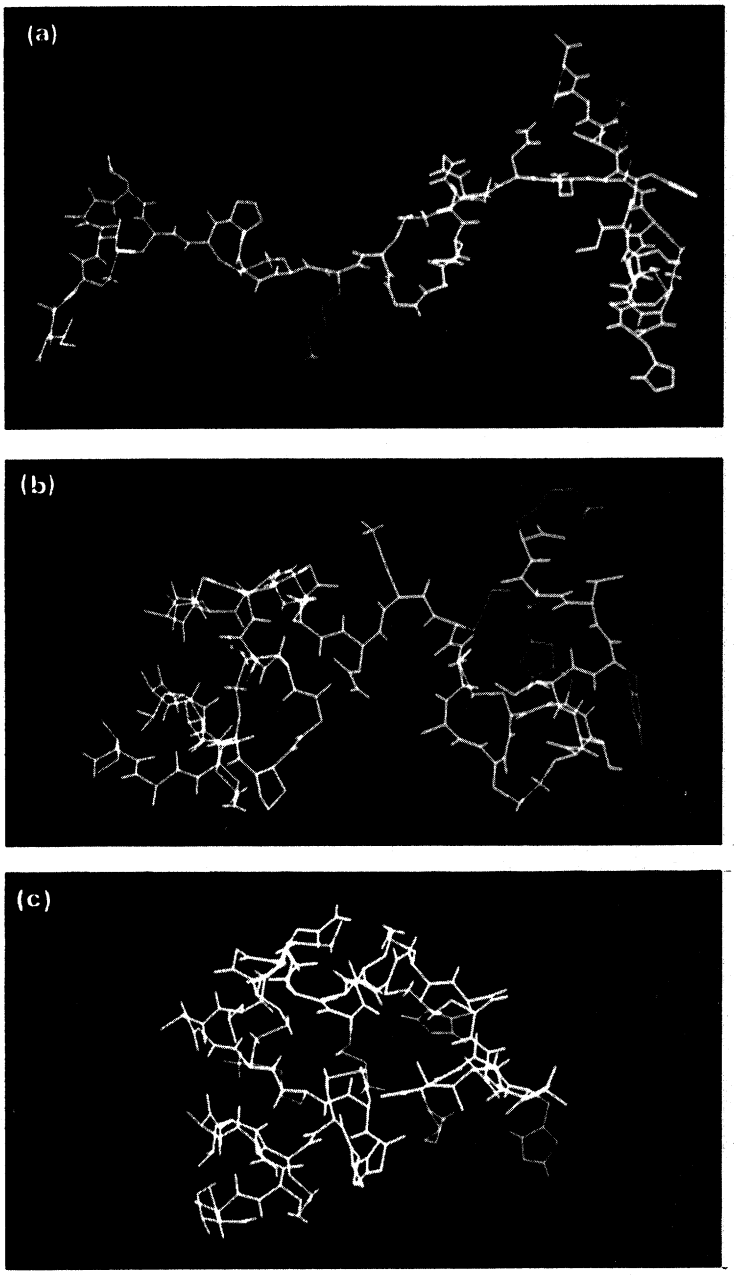


Fig. 3

Initial structure of nisin in a random-coil conformation. The so-called Ramachandran angles, ϕ (the αC_i–amino-N_i angle) and ψ (the αC_i–carbonyl-C_i angle) of the peptide bonds were initially set at –60° and 60°, respectively; the ω angle between carbonyl-C_i and amino-N_{i+1} atoms was kept constant at 180°. (a), the open structure (blue, the backbone and neutral side chains; purple, basic groups except histidine; green, hydrophobic regions; yellow, sulfur). (b), the disulfide cross-linked energy-minimized structure of nisin after simulated annealing for 40 picoseconds at 350 K. Colors are as in part (a), except histidine residues are purple. (c), energy-minimized structures of nisin subjected to complete simulated annealing at temperatures between 1000 K and 300 K for a total of 100 picoseconds.

Molecular dynamics and simulated annealing

The previous paragraph considered static structures only. However, the dynamic motion of molecules in solution contributes to their functionality. The molecular dynamics approach is a method of studying molecular conformation as a function of time. Each atom in the molecule is assigned a kinetic energy in the form of a velocity term, which can be related to local temperatures as well as to the average temperature of the system. Such calculations apply to molecules that are 'in vacuum' or in the presence of an appropriate number of molecules of a solvent such as water. In addition, the system being modeled can be studied at a constant temperature, volume or, in the future, pressure; in the case of pressure, a periodic boundary condition is defined to confine the system within a prescribed volume. For such molecular dynamics calculations, a force field describing the potential energy is combined with Newton's second law of motion:

$$F_i = m_i a_i = m_i \frac{dv_i}{dt} = m_i \frac{d^2 x_i}{dt^2} = -\nabla_i E \qquad (1)$$

where F_i is the force on atom i, which has mass m_i, velocity v_i, acceleration a_i and position x_i; ∇_i is the gradient or the derivative with respect to the position of atom i; t is time; E is the potential energy of the molecule described by the force field, and is a function of the positions of all atoms in the molecule. Equation 1 is integrated over various time intervals using a numerical integration method. The time intervals must be small (usually one femtosecond) compared with the time period associated with the highest frequency of motion within the molecule (usually that of stretching a bond associated with a hydrogen atom). Numerical integration of Eqn 1 over one-femtosecond intervals for 100 picoseconds for a protein molecule of ≥2000 atoms requires a fast computer with a large memory. Such calculations can model the motions of molecules in solution. Time-dependent geometric parameters can also be modeled; for example, the distance from the center of movement for a group of atoms may be related to correlation times derived from NMR, electron paramagnetic resonance or fluorescence spectroscopy experiments.

Composite sweep graphs of the dynamics, over 25 picoseconds, of a water molecule at 300K or 600K are presented in Fig. 2 to illustrate the results of molecular dynamics calculations. The graphs illustrate a combination of the rotational and angle-bending motions of the water molecule, and clearly show the increase in molecular motion with increasing temperature.

Molecular dynamics calculations also facilitate the process known as simulated annealing[16]. It is well documented that local minima exist in the potential energy surface of molecular conformations. Hence, energy-minimization techniques can give rise to erroneous results when a calculation becomes trapped within such barriers, and may yield a structure that is not in the lowest energy state. Molecular dynamics calculations can overcome such computational barriers and permit calculations to continue until fluctuation about the lowest energy state occurs. However, long computation times that are

unrealistic may be needed. The simulated annealing method can help to overcome this problem. In simulated annealing, molecular dynamics calculations are performed for high temperatures (1000 K) until a constant energy is reached. The system is then 'cooled down' and further molecular dynamics calculations are performed at lower temperatures; each calculation is continued until the total energy is constant over time. When a structure corresponding to ambient temperature (or lower) is obtained, its energy can be minimized to yield the 'best' (lowest-energy) model. The simulated annealing calculation samples a large fraction of conformational space and minimizes the problem of local energy minima.

Building a nisin template structure

In the modeling of polypeptides and proteins, an extremely large number of possible conformational states exist, and a reliable methodology for choosing the initial structure based on experimental evidence is essential when attempting to build a protein structure. One approach is to use sequence-based predictions of secondary structure[17] in conjunction with circular dichroism or FTIR spectroscopy. Although the programs that predict the secondary structure of a protein from its amino acid sequence are based on data from X-ray diffraction analysis, the various programs yield different secondary structures. Where a consensus occurs among the different prediction methods, more reliance can be placed upon the estimates. In the absence of such consensus, the programs must be used in conjunction with global secondary structure experiments to find one or more plausible initial structures.

The modeling of nisin, a simple polypeptide of 34 residues with 5 cross-linked lanthionine bonds, is presented to demonstrate a possible approach to building a protein structure using molecular modeling techniques (Fig. 3). Nisin is an antibiotic peptide, which can be used in some dairy products and which may soon be used as a preservative in other food systems; it has recently been expressed in *Escherichia coli*[18]. For this demonstration, cysteine residues were substituted for the sulfur cross-linked lanthionines, and the dehydroalanine and dehydrobutyrine residues were replaced by alanine. The sequence-based secondary structure predictions were ambiguous, and suggested an unordered structure. The peptide has not yet been characterized spectroscopically; thus, no experimental secondary structure results were available. As any definite secondary structure prediction was unavailable, a pseudo-random structure with $\phi = -60°$ and $\psi = 60°$ was chosen for the initial structure. The open structure is presented in Fig. 3a. Because nisin is normally used under acidic conditions, all histidine residues were protonated, which yields a single positive charge. Disulfide bonds were added to the appropriate residues, and the structure was 'energy minimized' by simulated annealing at 350 K for 40 picoseconds (Fig. 3b). This structure was then subjected to simulated annealing calculations to the lowest total energy at each of 1000 K, 500 K, 400 K and 300 K for a total of 100 picoseconds (Fig. 3c). Comparison of Figs 3a and 3c clearly shows how the simulated annealing process can overcome the problem of local energy minima.

Thus, a useful model is now available to determine surface residues that could be modified by genetic engineering to improve the properties of the molecule. The nisin structure (Fig. 3) appears to consist of two domains; it is suggested that one of the histidine residues may be in a small cleft and, therefore, have an altered pK_a reminiscent of some enzyme systems[19]. In subtilin[18], a similar peptide in which histidine is replaced by asparagine, a difference in the small cleft is observed. Thus, molecular modeling of antibiotic agents, their probable receptor sites and the interaction of the two molecules can suggest means of developing new and improved antibiotics by the introduction of point mutations.

Undoubtedly, as the technology of molecular modeling develops, new applications in food research will be found. An extension of this review, to be published in a future issue of *Trends in Food Science & Technology*, will examine potential applications in more detail, in particular the applications of molecular modeling to the study of the structure–function relationships of food proteins.

References

1 Richardson, T., Oh, S., Jiménez-Flores, R., Kumosinski, T.F., Brown, E.M. and Farrell, H.M., Jr in *Advanced Dairy Chemistry* (Fox, P.F., ed.), Elsevier (in press)
2 Cantor, C.R. and Schimmel, P.R. (1980) *Biophysical Chemistry Part II: Techniques for the Study of Biological Structure and Function*, pp. 687–791, W.H. Freeman
3 Wüthrich, K. (1989) *Science* 243, 45–50
4 Yang, J.T., Wu, C-S.C. and Martinez, H.M. (1986) *Methods Enzymol.* 130, 208–269
5 Susi, H. and Byler, D.M. (1986) *Methods Enzymol.* 130, 290–311
6 Krimm, S. and Bandekar, J. (1986) *Adv. Protein Chem.* 38, 181–364
7 Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) *J. Mol. Biol.* 112, 535–542
8 Clark, T. (1985) *A Handbook of Computational Chemistry: A Practical Guide to Chemical Structure and Energy Calculations*, John Wiley and Sons
9 Allinger, N.L. (1977) *J. Am. Chem. Soc.* 99, 8127–8134
10 Kalman, B.L. (1982) *Technical Memo No. 46*, Department of Computer Science, Washington University, MO, USA
11 Nemethy, G. and Scheraga, H.A. (1990) *FASEB J.* 4, 3189–3197
12 Nemethy, G., Pottle, M.S. and Scheraga, H.A. (1983) *J. Phys. Chem.* 87, 2361–2381
13 Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983) *J. Comput. Chem.* 4, 187–217
14 Weiner, S.J., Kollman, P.A., Nguyen, D.T. and Case, D.A. (1986) *J. Comput. Chem.* 7, 230–252
15 Papostephanou, C. and Frantz, M. (1978) in *Analytical Profiles of Drug Substances, Vol. 7* (Florey, K., ed.), pp. 401–422, Academic Press
16 Wilson, S.R. and Cui, W. (1990) *Biopolymers* 29, 225–235
17 Lin, T-H., Quinn, T.P., Grandgenett, D. and Walsh, M.T. (1989) *Proteins Struct. Funct. Genet.* 5, 156–165
18 Buchman, G.W., Banerjee, S. and Hansen, J.N. (1988) *J. Biol. Chem.* 263, 16260–16266
19 Timasheff, S.N. (1970) in *The Enzymes, Vol. II* (Boyer, P.D., ed.), pp. 430–440, Academic Press